Psychomusicology, 19, 15-40

THE IMPACT OF MUSIC ON SUBJECTIVE AND PHYSIOLOGICAL INDICES OF EMOTION WHILE VIEWING FILMS

Robert J. Ellis
Department of Psychology
The Ohio State University

Robert F. Simons
Department of Psychology
University of Delaware

Despite continued acknowledgment of the importance of music in viewing films, empirical studies investigating the interaction of music with film are conspicuously absent. Evidence from a few isolated cognitive studies suggests that the relationship between music and film is additive. Little is known, however, about the physiology of emotional response to viewing film–music stimuli. The present study utilized both self report as well as physiological indices to investigate the nature of the film and music relationship. Six-second films (having either negative or positive valence and low- or high-arousal) were paired with excerpts from instrumental classical music (pretested for valence and arousal). Results indicate a fairly straight-forward, additive relationship in terms of emotion self report. The modulating role of music on physiological reactions to film, however, was more complex. This study corroborates previous evidence regarding the subjective experience of viewing images with music. Physiological evidence, however, suggests that the interactions between music and film not always are predictable.

The scientific examination of the physiological correlates of human emotional response dates back to the later nineteenth century, with the work of Darwin (1872/1969; Lewis & Haviland-Jones, 2000). Since that time, many words have been written in an attempt to define what an emotion is—with varying degrees of success. Two general methods of classifying emotional response, however, have earned particular recognition: the *categorical* and *dimensional* approaches.

The categorical or *discrete emotions* approach (e.g., Izard, 1977) posits the existence of several "fundamental" emotions (e.g., happiness, sadness, anger, fear, disgust) that can function either independently or as members of an interrelated group (Neumann & Waldstein, 2001). A major criticism of this approach, however, is the lack of an agreed-upon set number of these fundamental emotions (Sloboda & Juslin, 2001). In various iterations of the theory, this number ranges anywhere from seven through more than a dozen (Izard, 1993).

Dimensional models of emotion date back over a century (Russell, 1997). Significant recent iterations include those by Russell (1980), Russell & Feldman-Barrett (1999), and Lang (1995). The dimensional model posits that every emotion can be expressed as a function of (usually) two "core" dimensions, frequently referred to as *valence* and *arousal*. Valence is described in terms of a range of response from positive (happiness, pleasure) to negative (sadness, displeasure). Arousal ranges from low (indicating calmness, peacefulness) to high (excitement, energy). These two dimensions can be aligned orthogonally so as to produce four quadrants (Figure 1), each of which represents one of four general valence–arousal categories: (1) positive valence, low-arousal (PV/LA); (2) positive valence, high-arousal (PV/HA); (3) negative valence, high-arousal (NV/HA); and (4) negative valence, low-arousal (NV/LA). Specific emotions can then be plotted as a function of the two dimensions; "anger" is NV/HA, for example, while "joy" is PV/HA. Statistical evidence appears to favor such a two-dimensional, valence–arousal approach to emotion (Faith and Thayer, 2001).

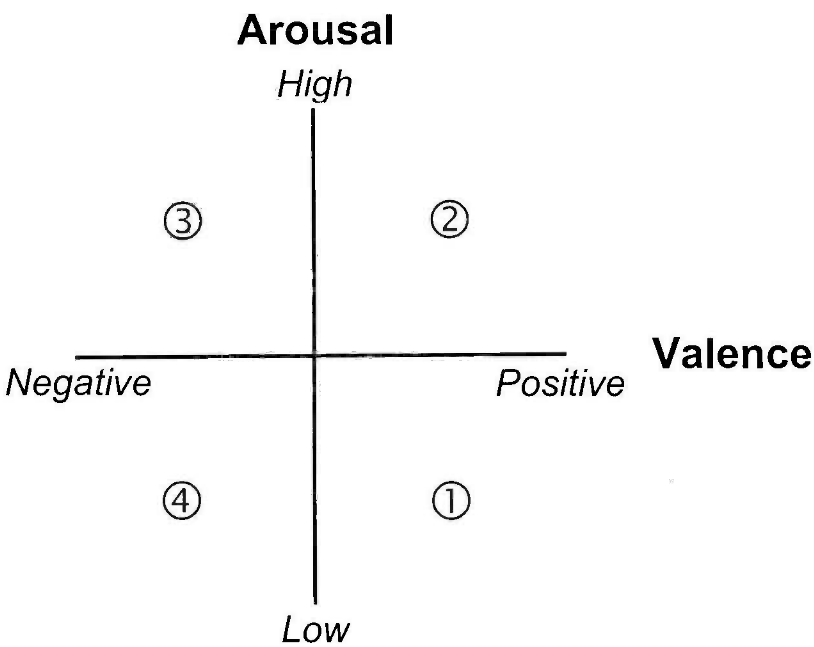

Figure 1. The dimensional model of emotion. Labels (1) through (4) indicate the four general valence–arousal categories as described in the text.

The dimensional approach is useful not only in classifying subjective (that is, verbal) labels of emotion, but also changes in physiology. Studies in which subjects view either still pictures (e.g., Lang, Greenwald, Bradley, & Hamm, 1993; Lang, Öhman, & Vaitl, 1988) or moving images (Detenber, 1995; Detenber, Simons, Detenber, Roedema, & Reiss, 1999; Simons, Roedema, & Reiss, 2000) suggest that the valence and arousal of the *content* of a visual image is consistently linked with changes in physiology—specifically, facial muscle activity, heart rate, and electrodermal activity.

Facial muscle (electromyographic, or EMG) activity is measured from two muscles: *zygomaticus major* (involved in the smile response) and *corrugator supercilli* (to contract and lower the brows, during frowning). Research has shown that EMG activity over the brow increases for *negatively*-valenced visual images, while activity over the cheek increases in response to viewing *positively*-valenced images (Bradley & Lang, 2000b; Cacioppo, Berntson, Larsen, Poehlmann, & Ito, 2000). Change in heart rate (HR) across the viewing period is also associated with the valence of a visual image: Negative images prompt more HR deceleration than positive images (Simons et al., 1999). Finally, electrodermal activity, here referred to as the skin conductance response (SCR), is a measure of palmar sweat gland activity and is closely linked with a subject's level of arousal. Visual images that are highly arousing elicit larger SCRs than images that are less arousing (Simons et al., 1999).

Music and Emotion: Physiological Evidence

Just as visual stimuli have been shown to induce physiological changes, research suggests that listening to music also can impact physiology. Perhaps the earliest scientific examination of this matter was conducted by the French composer Grétry (1741–1813), who attempted to measure the cardiovascular effects of singing by taking his pulse (Scherer & Zentner, 2001, p. 374). Interest in this topic has increased dramatically since that time. Bartlett's (1996) review references over 150 20th-century investigations of HR, SCR, respiratory, blood pressure, muscle tension, and biochemical responses to music. Other studies have endeavored to pair structural features of music to changes in physiology (Edwards, Eagle, Pennebaker, & Tunks, 1991). Neuroimaging techniques also have been useful in understanding the link between music and emotion (e.g., Blood & Zatorre, 2001).

Additional studies (Bolivar, Cohen, & Fentress, 1994; Boltz, 2001; Krumhansl, 1997; Lavy, 2001; Sloboda & Juslin, 2001) match structural aspects of music to a dimensional model of emotion, as summarized in Table 1. It is apparent that both visual and musical stimuli are capable of producing significant changes in levels of physiological activity, and that both can be examined within a dimensional framework of emotion. It is surprising and unfortunate, therefore, that there has been so little investigation of the physiology of emotional response when visual and musical stimuli are presented *simultaneously*. This lack of research, however, only serves to mirror the rather unfortunate position that the scientific study of music and emotion has occupied within a more mainstream literature. Its place within the field of music psychology is tenuous at best, even today (Juslin & Sloboda, 2001; Juslin & Zetner, 2002). Deutsch's (1999) thorough survey on the psychology of music does not discuss emotion; similarly, a recent scholarly text on emotion (Lewis & Haviland-Jones, 2000) makes no reference to music.

Table 1

The Relationship Between Musical Structure and Perceived Emotion

Structural features	Rated emotion
Major key, consonant harmonies	Positive valence
Minor key, dissonant harmonies	Negative valence
Slow tempo, regular rhythm or meter	Low arousal
Fast tempo, irregular rhythm	High arousal

Music and Film: An Introduction

Music's role in film dates back almost to the beginning of the film medium itself; today, a film's musical score is so ubiquitous that its absence almost would seem cause for alarm. Despite the seeming prominence of music during film presentations, however, it is generally conceded that the systematic examination of film music has been roundly neglected by film theory (Donnelly, 2001; Tan, 1996), musicology (Prendergast, 1992), and music psychology (Cohen, 2001).

Of the relatively few studies that have investigated music and film, nearly all have utilized exclusively cognitive or *schema-based* approaches (e.g., Bolivar, Cohen, & Fentress, 1994; Bullerjahn & Güldenring, 1994; Sirius & Clark, 1994) without including physiological indices. To wit, Boltz (2001) cites only one study (Thayer & Levenson, 1983) that has investigated physiological responses to compound film–music stimuli.

One common finding reported by the studies mentioned above is that music adds to the emotional impact of film. Bolivar et al. (1994) paired films (depicting a pair of wolves engaged in either friendly or aggressive behavior) with music (pretested as sounding "friendly" or "aggressive"). Music had a main effect on the judgments of the scenes: Both friendly and aggressive interactions received higher ratings of friendliness when shown with friendly (versus aggressive) music. Sirius and Clark (1994) paired a film of moving geometric shapes (small or large, fast or slow) with excerpts of music. Subjects were asked to provide ratings along dimensions of "activity" and "potency." The authors found that both fast and slow moving shapes were judged as *active* when shown with highly active music, and both small and large shapes as *more active* when shown with high potency music.

The results of Thayer and Levenson (1983) suggest that this additive relationship between music and film also is present within physiological indices of emotion. In their study, musical scores were added to the (verbal) soundtrack of an industrial safety film depicting three accident scenes. Three conditions were created: (a) the film accompanied with no music, (b) the film accompanied by a "horror" musical score, and (c) the film accompanied by a documentary musical score. Although several physiological variables were recorded, results indicated that only SCR differentiated the three film score conditions, with the highest average level of SCR activity occurring in response to the horror music score, followed respectively by the control (no music) condition and documentary music score. Thus the overall relationship appeared to be additive: horror (i.e., high-arousal) music made the experience of watching the film more arousing overall—and documentary (i.e., low-arousal) music less arousing—than when no music was present.

More recently, Ellis and Simons (2002) reported that presenting a film with its associated soundtrack (versus with the volume muted) impacted both subjective and physiological indices of valence and arousal. Over the course of the experiment, subjects viewed a collection of short films (6 s) twice: once with their associated soundtracks (e.g., background noise, nature sounds, instrumental music) and once without. A control group listened to the soundtracks without viewing the films. The results of this study suggested that responses to the audiovisual stimuli were consistent with a superimposition of the visual and auditory stimuli presented independently. In other words, the response to the compound audiovisual stimulus appeared similar to a simple addition of the response to the two constituents (i.e., response to film plus response to soundtrack). This effect was present in both self-report and physiological indices of arousal and, to a lesser degree, valence. The study made no effort, however, to determine whether different types of soundtracks (e.g., environmental sounds, animal noises, instrumental music) might have impacted self-report and physiological indices differently. This is of critical importance, as recent evidence (Gomez, Ott, & Danuser, 2002) suggests that music and sound impact physiology in different ways, due perhaps to the degree in which the acoustic stimuli activated visual imagery in the subjects. Thus, the present study will endeavor to improve upon the mixed results of both Thayer and Levenson (1983) and Ellis and Simons (2002).

The Present Study

The present study was designed to determine whether the relationship between simultaneously presented music and film is additive when indexed by the self-report of valence and arousal, and to examine the extent to which this relationship is reflected in physiological indices of emotional response.

Our expectations, thus, were as follows: (a) Positive films would be viewed more positively (evidenced by more positive ratings of self reported valence, greater zygomatic activity, and less heart rate deceleration) than negative films; (b) high-arousal films would be more arousing (evidenced by higher subjective ratings of arousal and greater skin conductance activity) than low-arousal films; (c) films shown with positive music would be viewed more positively (again evidenced by more positive ratings of self reported valence, greater zygomatic activity, and less heart rate deceleration) than films shown with negative music; and (d) films shown with high-arousal music would be more arousing (evidenced by higher subjective ratings of arousal and greater skin conductance activity) than films shown with low-arousal music.

Four hypotheses emerge out of these basic predictions. The first is that film clips can be reliably scaled on the two-dimensional emotional spectrum of valence and arousal; the second is that these emotion dimensions would be associated with specific physiological sequelae. The third and forth posit that the relationship between simultaneously presented music and film would be additive in both self-report of emotion and its concomitant physiology, respectively.

Method

Subjects

Thirty-six undergraduate students at the University of Delaware participated in this study in partial fulfillment of the research component of their general psychology class. Selection was based on two criteria: (a) that students rated their knowledge of classical music as either "none" or "slight," and (b) that they possessed less than one year of formal music training. These volunteers were selected randomly from a larger pool of 1,620 students enrolled in general psychology in the semester in which the research was conducted. The majority of the pool was enrolled in the College of Arts and Sciences (36%), though all colleges were represented in the pool. The vast majority of the students were freshmen (79%), and the declared major areas of study numbered 84. Although the course content was general psychology, only 6% of the students were psychology majors. In short, the students, though primarily freshman, were in all other ways very representative of the student body as a whole. Of the original 36 subjects, data from 2 subjects were eliminated due to technical problems during data collection. The final sample of 34 contained 17 men and 17 women.

Stimuli

Visual. The stimuli consisted of thirty-two 6 s films, selected from a large collection of excerpts from film and television programs that previously had been standardized by Detenber (1995). The present subset contained eight films in each of the four valence–arousal categories defined in Figure 1, namely, NV/LA, NV/HA, PV/LA, and PV/HA. The films were selected and categorized on the basis of prior ratings of both valence and arousal (Detenber, 1995). Table 2 shows the average valence and arousal ratings and standard deviations for each category.

Table 2

Mean (Pre-Sort) Valence and Arousal Ratings for the Four Valence–Arousal Categories of Films, Based on Data from Detenber (1995)

Film Category (Valence/Arousal)	Rated Valence		Rated Arousal	
	Value	*SD*	Value	*SD*
Negative/Low	3.5	0.5	4.2	0.4
Negative/High	2.2	0.3	7.1	0.3
Positive/Low	7.1	0.3	3.8	0.2
Positive/High	7.1	0.3	6.6	0.5

Audio. Eight 4.5 min excerpts of instrumental music (hereafter referred to as "tracks") were selected, so as to have two tracks in each of the four valence–arousal categories. For more detailed information about the music selected, see the Appendix. The length of each track was similar to those used in other studies investigating emotional response to music (e.g., Boltz, 2001; Faith & Thayer, 2001; Krumhansl, 1997; Lavy, 2001). Tracks were placed *a priori* into one of the four valence–arousal categories based on structural characteristics (tempo, mode, loudness, harmonic movement) known to influence judgments of perceived emotion (cf. Table 1), and these assignments were confirmed by six pilot subjects who were asked to rate each track in terms of both valence and arousal.

Combined audiovisual. Eight "clips" then were created using Adobe *Premiere 6.5* software, in the following manner. One track, heard in its entirety, was overlaid with eight different films (two each from the four valence–arousal categories), shown at 30 s intervals. Figure 2 illustrates the design of a representative clip. Each of the 32 films was presented twice over the course of the experiment, paired with tracks from *opposing* valence–arousal categories. For example, if a film was shown first with PV/LA music, it was shown a second time with NV/HA music; or if a film was first paired with NV/LA music, it was shown again with PV/HA. In this manner, over the course of the experiment every film was presented with both negatively- and positively-valenced music, and both low- and high-arousal music. Since the content of a given film was identical over the two viewings, any differences in subjective or physiological responses between the two viewings could thus be attributed to the effects of music on the audiovisual experience.

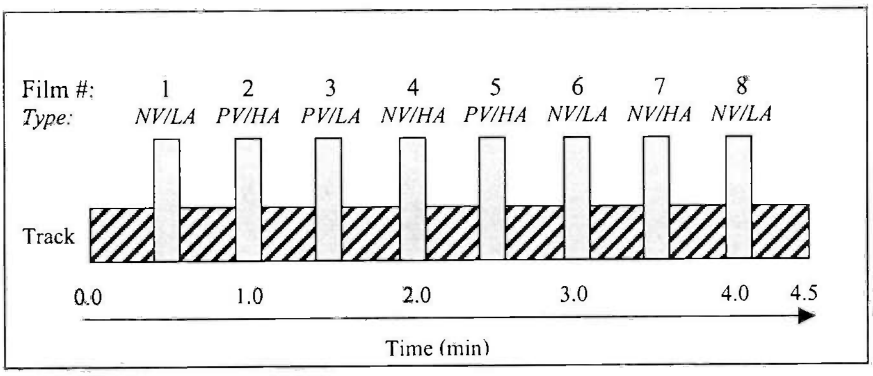

Figure 2. Construction of a sample clip. Each track was 4.5 min in duration and played without interruption. At 30 s intervals, a 6 s film was presented on screen. A 30 s baseline preceded the presentation of the first film presentation. *Type* refers to the valence–arousal category.

Response Measurement

Self report. Subjects rated their emotional reactions to each stimulus on the dimensions of valence (negative to positive) and arousal (low to high) via a nine-point, pencil-and-paper version of Lang's Self-Assessment Manikin (SAM, Lang, 1980). The respondent marked either on or between two of the five manikins on each scale, indicating the direction or magnitude of their emotional response. For valence, the manikin was shown with facial expressions ranging from a broad smile (most positive; later quantified as a 9) to a severe frown (most negative; later quantified as a 1). Arousal was similarly represented on a second scale by five graphics depicting ever increasing visceral agitation, from low (1) to high (9). A response was made after each film presentation (a total of 64 responses). High SAM valence scores thus indicate pleasure, and high SAM arousal scores excitement.

Physiological recording. The recording and measurement of physiological data were similar to those of Simons et al. (1999). Skin conductance responses were recorded using a locally constructed constant voltage (0.5 V) circuit in conjunction with a Grass Model 7D polygraph. Beckman Standard (0.5 cm^2) Ag/AgCl electrodes were placed on the thenar and hypothenar eminences of the palm on the subject's left hand, and Johnson & Johnson K-Y Jelly was used as the electrolyte.

EMG recordings from the face were obtained by placing Med-Associates miniature Ag/AgCl electrodes over the subject's left zygomatic and corrugator muscles. The area of skin under each pair of electrodes was abraded until impedence was below 10K Ohms, and Hewlett-Packard Redux paste was used as the electrolyte. The raw EMG (bandpass = 3 to 500 Hz) was full wave rectified and integrated (time constant = 50 ms) using a Grass Model 7P3 Wideband Amplifier/Integrator.

Heart rate was obtained by attaching a Grass Photoelectric Transducer Model PPS to the subject's right ear lobe. The photocell output was fed into a Grass Model 7P1 Low Level DC Preamplifier and Model 7D Driver Amplifier (bandpass = 1.6 to 3.0 Hz) and then into a locally-constructed Schmitt trigger which shaped the pulse prior to computer detection.

Procedure

Subjects were run individually. While in the control room, each subject was provided with a short verbal description regarding the nature of the study, the stimuli, and the recording techniques before reading and signing an informed consent form. EMG and skin conductance electrodes were attached, and the subject was led into an adjacent room and seated in a comfortable armchair, approximately 1.4 m in front of a large projection screen on which the films were to be presented. Sound was delivered in stereo via two speakers on the floor, located directly below the screen and approximately one meter apart. Once seated, the photocell was attached to the subject's left ear, and the quality of the physiological recordings assessed from the control room.

Subjects then were instructed regarding the task and ratings procedure. They were informed that they would be viewing short films while listening to music. They were told that the music would run throughout the course of the experiment, changing periodically, and that they should "treat the music as they would music during a movie." Subjects were told that after the conclusion of each film segment, they should use the booklet to indicate their emotional state while watching (N.B.: The difference between emotion "perceived" and emotion "felt" is discussed later in this article). Subjects were shown a neutral practice trial to familiarize them with the nature and timing of the stimuli, as well as with using the SAM. The practice clip was a short (1.5 min) excerpt of music during which two films were presented. The experiment commenced if the instructions were understood and properly followed during the practice trials, and if the physiological data were free of noise and artifact.

The experiment proper consisted of 64 trials (eight trials per clip) under the control of two linked laboratory PCs—one to control the presentation of clips using *Presentation* software (Neurobehavioral Systems, Inc.) and the other to digitize and store the physiological data (*VPM* software; Cook, 1999) for later offline analysis. The clips were presented in one of two pseudorandom orders, and the valence-arousal category of any two successive clips was always different. For each of the 64 film presentations, collection of physiological data commenced 2 s prior to the on-screen onset of the film and terminated 2 s after the image disappeared (10 s total). After each film presentation, an on-screen text message indicated that subjects should make their SAM valence and arousal ratings.

Halfway through the experiment (after the conclusion of the fourth clip), the experimenter reentered the viewing room to provide a short break and to ensure that subjects were on the appropriate page of the ratings booklet. At the conclusion of the experiment, subjects were given a debriefing form and thanked for their participation.

Data Reduction

Data reduction and analysis procedures were similar to those of Simons et al. (1999). The skin conductance and the two facial EMG channels were digitized at 50 samples per second. Discrete skin conductance responses (SCRs) that began with an onset latency of 0.5 to 4 s following the onset of each film were displayed graphically, trial by trial, and quantified by visual inspection. Magnitude of the SCR was defined as the difference, in μSiemens, between the identified peak and onset points. SCR data were subjected to a natural log transformation prior to statistical analysis.

Corrugator and zygomatic muscle activity for each trial were expressed as the difference between the mean value during the 6 s viewing period and the prestimulus mean. As the magnitude of response is much greater in the zygomatic muscle, data from both zygomatic and corrugator muscles were standardized within each respondent. Then a "pattern" score was computed for each trial by subtracting the zygomatic change from the corrugator change. Thus, a positive pattern score would reflect greater corrugator activity (more brow furrowing than smiling) and a *negative* emotional reaction, whereas a negative pattern score would reflect the opposite (more smiling than brow furrowing) and a *positive* emotional reaction.

The photocell recorded interbeat intervals that were converted to HR in beats per minute (BPM) per real time epoch (each epoch was one-quarter second). When epochs contained portions of two beats, each BPM rate was weighted according to the fraction of the epoch that it occupied. Heart rate waveforms were obtained from the edited HR record by averaging successive epochs, and deviating half-second averages during the 7 s post onset epoch from the 0.5 s average immediately preceding stimulus onset. Fourteen half-second averages, along with the onset point, constituted the HR data that then were submitted to statistical analysis.

Data Analysis

The SAM valence and arousal data and two of the dependent measures (EMG, SCR) were analyzed using a 2 x 2 x 2 x 2 x 2 repeated-measures ANOVA model. There were four independent variables, each with two levels: (a) Film Valence (negative or positive), (b) Film Arousal (low or high), (c) Music Valence (negative or positive), and (d) Music Arousal (low or high). The fifth variable, Gender, was a dependent measure. HR data were analyzed using a 2 x 2 x 2 x 2 x 2 x 15 repeated-measures ANOVA model. The first five variables were identical to the five variables just described. The sixth variable was time, with fifteen levels representing successive half-second averages obtained from HR data reduction.

Results

SAM Ratings

Emotion self report was measured using the Self-Assessment Manikin (SAM; Lang, 1980). The top panel of Figure 3 presents valence ratings as a function of film valence and music valence. Confirming that our selection of clips induced the anticipated emotion experience, film valence had a significant effect on reported levels of valence, $F(1,33) = 495.31, p < .001$, with positive films rated more positively (higher SAM valence scores) than negative films. Music valence also had a significant effect on perceived valence. Films shown with positive music were rated more positively than the same films shown with negative music, $F(1,33) = 42.49, p < .001$. There was no main effect for gender ($F < 1$), nor did gender interact with either film or music valence. The bottom panel of Figure 3 presents SAM arousal ratings as a function of film arousal and music arousal. Again, as expected based on our selection of film stimuli, high-arousal films were rated as more arousing (higher SAM arousal scores) than low-arousal films overall, $F(1,33) = 202.97, p < .001$. The arousal category of music also had a significant effect on rated arousal. Films shown with high-arousal music were judged as more arousing than films shown with low-arousal music, $F(1,33) = 19.16, p < .001$. Once again, there was neither a main effect for gender nor interactions between the gender and arousal variables.

One interaction that was present, however, was between film valence and film arousal, $F(1,33) = 121.68, p < .001$, as shown in Figure 4 (top). Negative valence, high-arousal (NV/HA) films were rated much more negatively than negative valence, low-arousal (NV/LA) films, while positive valence, high-arousal (PV/HA) films were rated somewhat more positively than positive valence, low-arousal (PV/LA) films. To further illuminate the role of arousal in negative versus positive films, a paired samples *t* test was conducted. Film arousal significantly differentiated SAM valence to a greater degree within negative films ($t[32] = 12.09, p < .001$) than positive films ($t[32] = 3.04, p < .005$).

Although music arousal did not have a main effect on rated valence ($p > .3$), it did interact with music valence, $F(1,33) = 18.19, p < .001$, as shown in the bottom panel of Figure 4. Namely, films shown with NV/HA music were rated more negatively than films shown with NV/LA music overall, whereas films shown with PV/HA music were rated more positively than films shown with PV/LA music overall. Once again, paired samples *t* tests indicated that music arousal significantly differentiated SAM valence within both negative ($t[32] = 3.07, p < .005$) and positive ($t[32] = 3.74, p < .005$) films. This interaction between music valence and music arousal is similar to the interaction between film valence and arousal reported above (Figure 4, top panel). Based on a comparison of their respective *F* ratios, however, it was apparent that the strength of interaction between film valence and film arousal was more substantial than music valence and music arousal. Also, *t* tests indicated that, while the relationship between SAM valence ratings to PV/LA and PV/HA

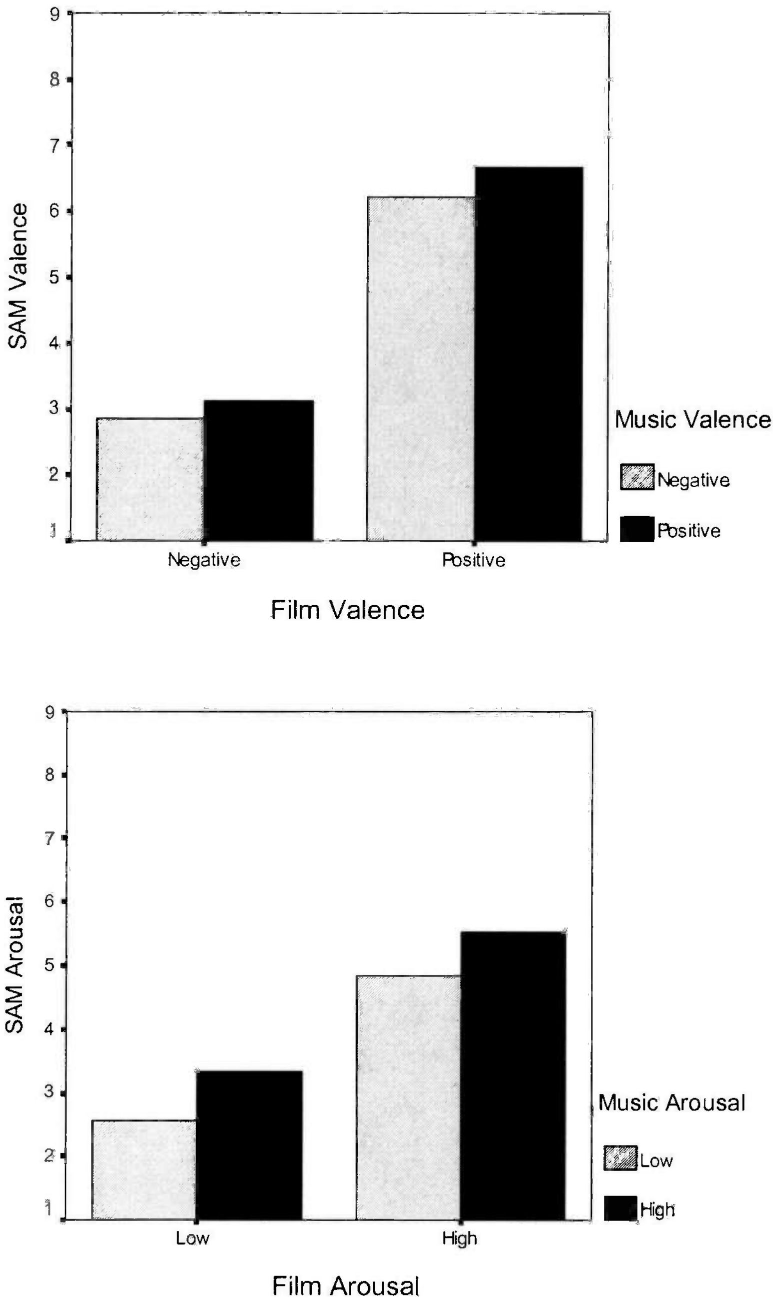

Figure 3. SAM valence ratings as a function of film valence and music valence (top) and SAM arousal ratings as a function of film arousal and music arousal (bottom).

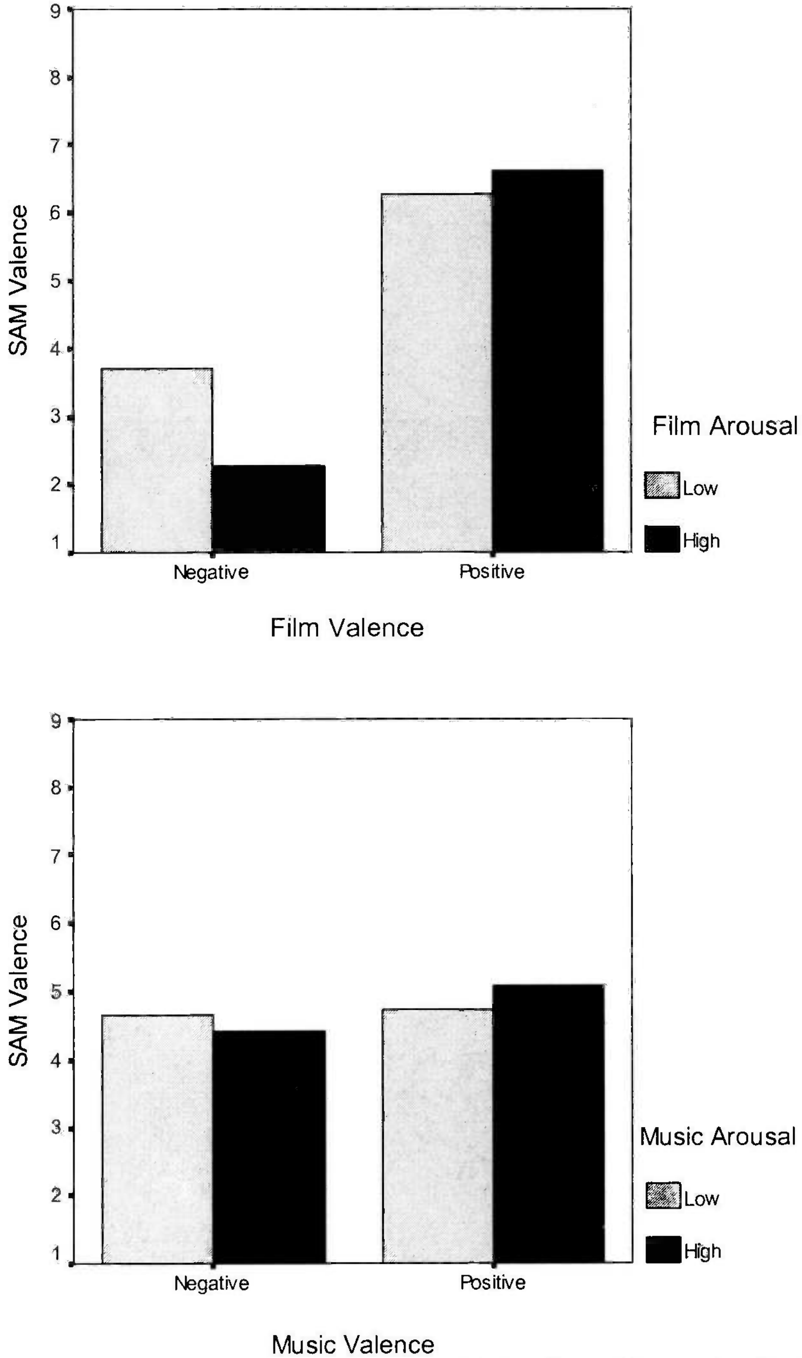

Figure 4. SAM valence ratings as a function of film valence and film arousal (top) and SAM valence ratings as a function of music valence and music arousal (bottom).

films (the right two bars in the top panel of Figure 4) compared to PV/LA and PV/HA music (the right two bars in the bottom panel of Fig 4) was similar, the difference in valence ratings between NV/HA and NV/LA films (the left two bars in Figure 4 top panel) was much more pronounced than the difference between NV/HA and NV/LA music (the left two bars in Figure 4 bottom panel), with NV/HA films rated much more negatively than NV/HA music.

Facial EMG

Figure 5 illustrates the relationship between the facial EMG pattern score and film valence, and the EMG pattern score and music valence. The pattern score varied with film valence, $F(1,33) = 18.86, p < .001$: Positive films prompted strong zygomatic responses (smile pattern) while negative films elicited strong corrugator responses (frown pattern). Although the figure suggests that music and film valence may have had an interactive effect on the EMG pattern score, the interaction did not approach statistical significance ($F[1,33] = 1.23, p > .25$). Neither film arousal ($F < 1$) nor music arousal ($F < 1$) had a main effect on the EMG pattern score. As with SAM data, there was neither a main effect for gender nor any interactions involving gender.

Heart Rate

The HR response to film stimuli was deceleratory, beginning shortly after stimulus presentation and remaining below baseline for the duration of the recording interval. This is consistent with our previous studies using similar film clips. Figure 6 presents half-second by half-second data as a function of film valence during musical excerpts pretested as negative (top panel) and positive (bottom panel). As expected, film valence had a significant main effect on heart rate deceleration, $F(1,33) = 4.26, p < .05$, with greater deceleration in response to viewing negative versus positive films. The development of the differentiation of response to negative versus positive films over time (Film Valence x Time) was reflected in both linear ($F_{lin}[1,33] = 4.22, p < .05$) and quadratic ($F_{quad}[1,33] = 9.40, p < .005$) interactions. There were no significant effects of film arousal on heart rate change. As Figure 6 illustrates, the effect of music on HR was to accentuate these differences. That is, when the film clip was negative, music had little impact, but there was an apparent synergy between positive films and positive music that did modulate the evoked HR response. Statistically, this took the form of a significant Film Valence x Music Valence x Time interaction ($F_{quad}[1,33] = 5.90, p < .05$). HR response was not significantly different between genders ($F < 1$).

Skin Conductance

Figure 7 (Part 1)shows SCR magnitude as a function of arousal (top panel) and valence (bottom panel) categories. As expected, high-arousal films evoked greater skin conductance responses than low-arousal films, $F(1,33) = 28.06, p < .001$. The arousal category of the music had a marginal effect on skin conductance, $F(1,33) = 3.67, p < .10$, see left panel. As shown in the top panel, larger SCRs were elicited in response to films when accompanied by high-

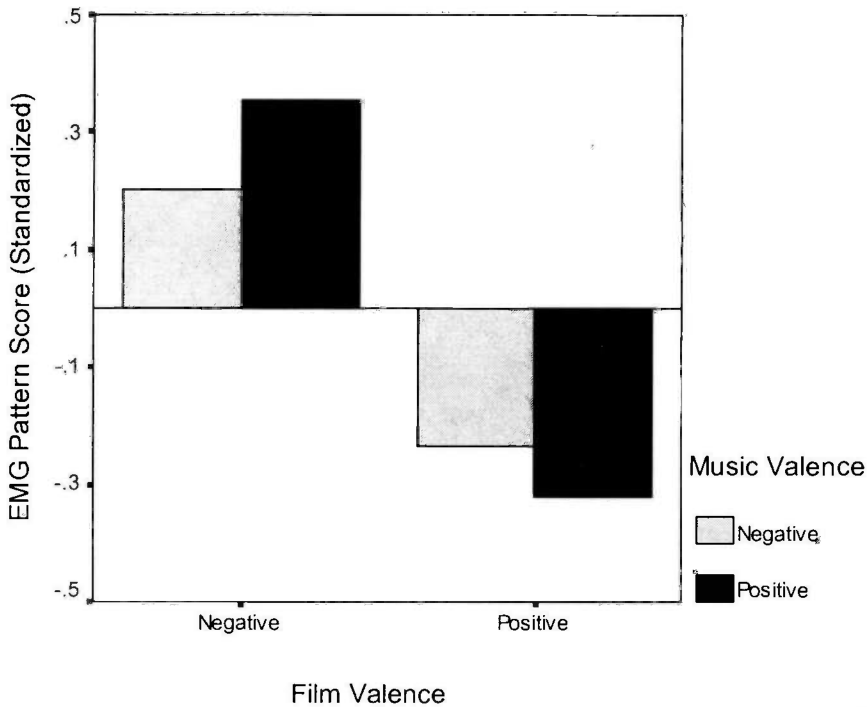

Figure 5. Facial EMG pattern score as a function of film valence and music valence.

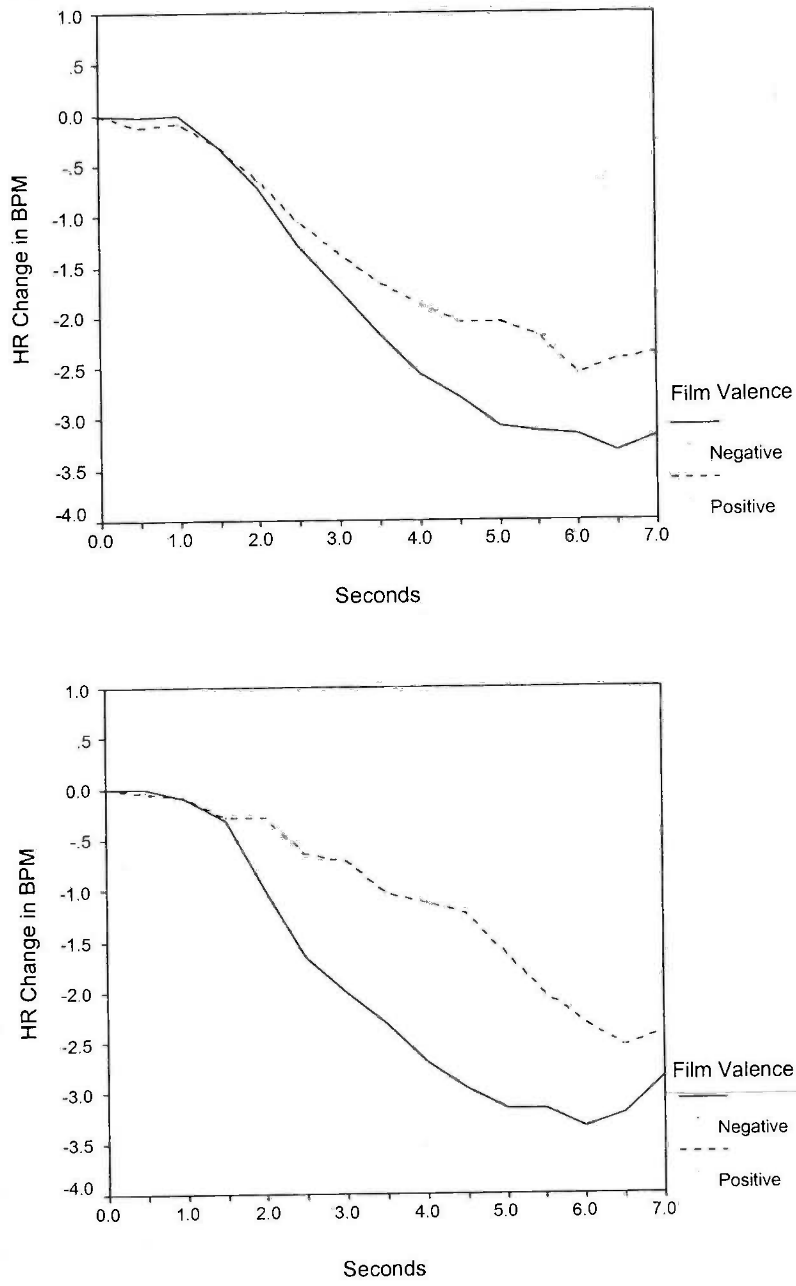

Figure 6. HR change as a function of film valence when accompanied by negative (top) and positive (bottom) music.

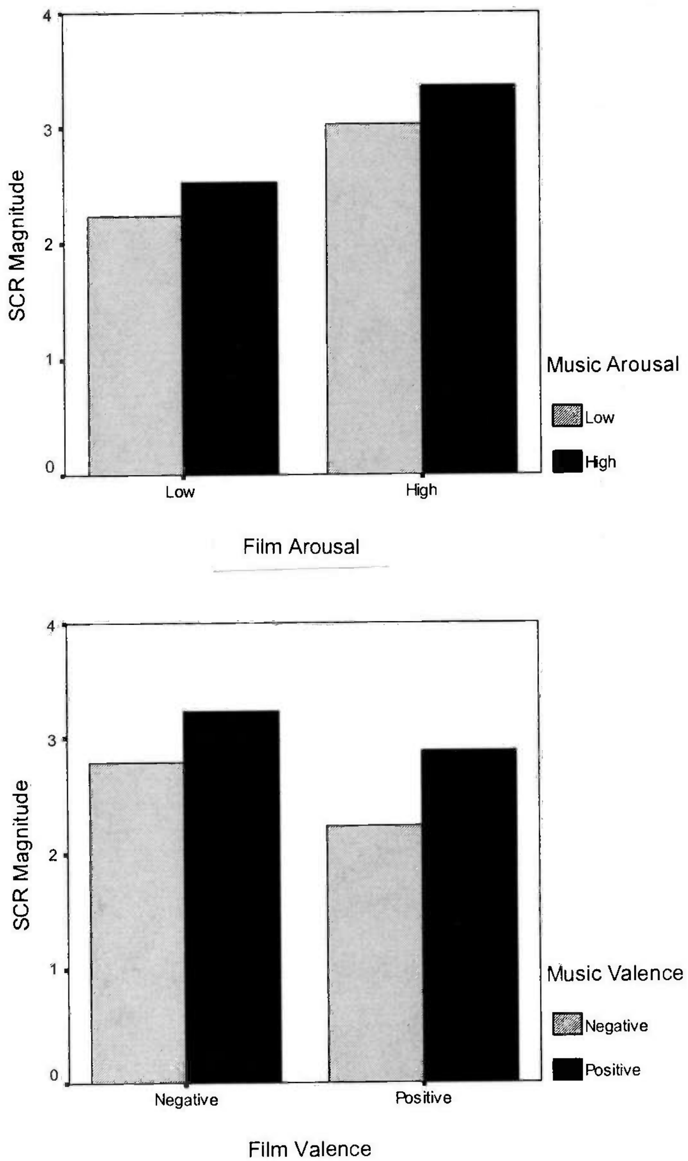

Figure 7 (Part 1). SCR magnitude as a function of arousal (top panel) and valence (bottom panel).

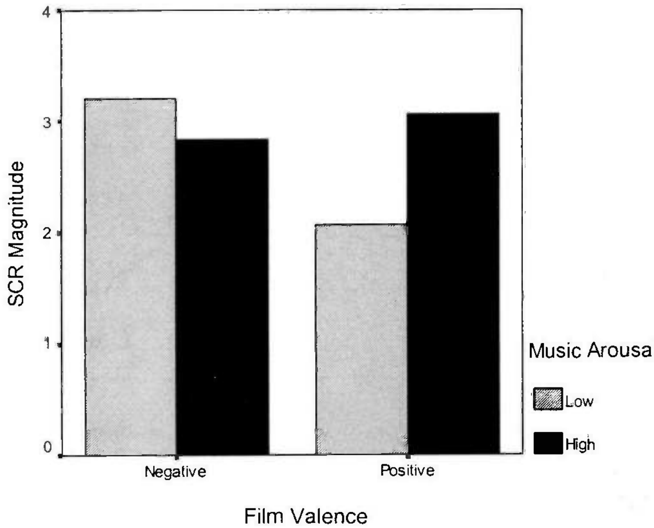

Figure 7 (Part 2). SCR magnitude as a function of film valence and music arousal.

arousal music. Both film valence ($F[1,33] = 5.66, p < .05$) and music valence ($F[1,33] = 7.48, p < .01$) had main effects on SCR, see bottom panel (Figure 7, Part 1). Negative films elicited larger SCRs than positive films overall, and films shown with positive music evoked larger SCRs than films shown with negative music overall.

One notable interaction occurred between film valence and music arousal, $F(1,33) = 14.31, p < .001$, as shown in Figure 7 (Part 2). A paired samples t test was conducted to see if music arousal modulated SCR magnitude during both negative versus positive films. This was not the case. Music arousal did not significantly affect SCR when the valence of the film was negative ($t[32] = 1.43, p > .1$); it did, however, when film valence was positive ($t[32] = 4.37, p < .001$). That is, high-arousal music elicited greater skin conductance activity than low-arousal music *only* when it was paired with positive films.

Discussion

The first two of the hypotheses tested in the current study emerged from previous findings from related studies (e.g., Lang et al., 1993; Simons et al., 1999); namely, that pictures—or in this case, films—could be scaled reliably on the two emotion dimensions of valence and arousal and that these emotion dimensions would be associated with specific physiological sequelae. As in our previous studies, these hypotheses received strong support. Although the majority of the films employed in the present study had not been used in previous studies (e.g., Ellis and Simons, 2002), they nonetheless elicited appropriate subjective reports of emotion experience, and these self reports of emotional valence and arousal were associated with the anticipated changes in peripheral physiology. That is, negatively-valenced films prompted greater HR deceleration and a corrugator dominant facial EMG pattern while positively-valenced films were associated with a zygomatic dominant facial pattern and an attenuated decelerative HR response. High-arousal films, regardless of their valence, gave rise to larger electrodermal responses than low-arousal films. Thus once again, specific physiological responses were associated with emotion states when emotions were scaled in terms of their valence and arousal properties.

The third and fourth hypotheses central to this study were directly concerned with the notion that music would modulate emotion states in a manner similar to the films, and that the relationship between simultaneously presented music and film would be additive in nature. The third hypothesis stated that this relationship would be apparent when emotion experience was self-assessed; and the fourth that whether this relationship would be apparent when physiological indices of emotion were evaluated. The third hypothesis was supported: Both music valence and music arousal had main effects on subjects' ratings. Positive music elicited higher SAM valence ratings of both positive and negative films, and high-arousal music elicited higher ratings of SAM arousal for both high- and low-arousal films.

The fourth hypothesis—that this additivity would extend to physiological responses—received mixed support. Surprisingly, differentiation of response to music valence was not at present in the valence-sensitive facial EMG. It was, however, evident in HR data, but only when film content was positive. The skin conductance response, normally associated with arousal, was in fact affected by music valence (larger SCRs when the music was positive), but like HR, only when the music was paired with positive films. Finally, while high-arousal music prompted larger SCRs, it did so only when the valence of the films was positive.

The lack of any differentiation between subjects' emotional responses based on their gender was consistent across both self report and physiology. Such a finding agrees with similar studies investigating subjective and physiological responses to viewing films (e.g., Bradley & Lang, 2000a; Detenber et al., 2000; Simons et al., 1999; Thayer & Levenson, 1983).

In summary, the anticipated main effects for music valence and music arousal that appeared within self-report indices of emotion were not always seen in the concomitant physiology, and the effects of music valence and music arousal found in the physiology did not always find correspondence in the self-report. The effect of music on emotional responses to viewing films, therefore, is complicated and not necessarily predictable.

The Visual Negativity Bias

As mentioned above, data from both HR and SCR reflected an interaction between film valence and music. Specifically, the effects of music arousal on skin conductance were more pronounced when the valence of the film was positive rather than negative; similarly, the effects of music valence on HR were present only when thevalence of the film was positive. Although these findings were somewhat unexpected, they are not without precedent.

As noted earlier, Bolivar et al. (1994) found that films depicting both friendly and aggressive social interactions among a pair of wolves were given higher ratings of friendliness when paired with friendly (versus aggressive) music. Significantly, however, the difference in ratings between the two musical conditions was smaller within aggressive films than within friendly films; aggressive exchanges were less influenced by friendly music than were friendly exchanges. The authors attributed this finding to the strong salience of the negative visual content, suggesting that music had less of an effect on subjective ratings of emotional response because the valence of the visual is highly negative. They cited a similar conclusion by Cohen (1993), who found that two contrasting musical soundtracks (hostile versus playful) differentiated valence ratings to a lesser degree when the visual content was violent versus when it was neutral.

Thayer and Levenson (1983) also suggested that music may be less able to impact the physiology of emotional response to viewing films if the content of the film is highly negative. Subjects viewed a film depicting industrial accident scenes with either horror (i.e., high-arousal) music, documentary (i.e., low-arousal) music, or no music. Although skin conductance activity across all music conditions was maximal during the climax of each accident scene, SCR was also found to interact with the musical condition. While there was a significant difference in peak SCR magnitude between the horror music and no music conditions for two of the three accident scenes, there was no significant difference in peak SCR levels between the documentary music and control conditions. That is, low-arousal music was unable to make a strongly-negative (and highly-arousing) film less arousing.

The Power of Film?

The present findings highlight what appears to be the primacy of *visual*—rather than auditory—content in shaping emotional response when both modalities are providing information. Both of the hypotheses regarding film valence and arousal were supported. Not only were main effects for subjective ratings of film valence and film arousal present, but they were larger than those for music valence and music arousal. The dramatic difference in self-reported valence in response to viewing NV/HA films (Figure 4) and hearing NV/HA music (Figure 3) suggests that films are better able to encapsulate an emotional state than is music. Additionally, physiological evidence was also more consistent with previous studies that used a similar stimulus set (Detenber et al., 2000; Ellis & Simons, 2002; Simons et al., 1999). A main effect for film valence was reflected in both EMG and HR data (positive films prompting more zygomatic activity and less HR deceleration than negative films), and a main effect for film arousal was reflected in SCR data, with greater activity in response to high-arousal films. Main effects for music valence and arousal were absent among physiological indices, as noted previously.

A "Binding Relationship" Between Music and Film

Given such strong evidence, it might seem appropriate to conclude that visual stimuli are more emotionally "affective" than musical stimuli when presented simultaneously. A different explanation, however, would suggest that it might not be a question of the music itself, but rather how music becomes incorporated into the narrative of the visual domain (Cohen, 2001). If film music's ability to evoke emotion is contingent upon its ability to bind with a visual narrative, and if the music were somehow prevented from associating with this narrative, the music may be seen as merely adjunctive rather than integral to the subsequent understanding and coloring of the overall emotional response.

An ideal binding situation would exist when an acoustic stimulus (whether musical or otherwise) could logically be incorporated into the visual diegesis (that is, the visual world) of a film; for example, viewing a bird and hearing chirping noises. If we were to see a bird and hear bird-like noises, we would logically conclude that the source of those noises *is* the bird. Music, however, is harder to incorporate into the visual diegesis of a film. Naturally, if in a film we saw an on-screen orchestra accompanying the orchestra playing in the soundtrack, the music could logically bind to the visual diegesis. In most feature films, however, this is rarely the case: Music is "non-diegetic," existing without any obvious source.

The present study further complicates this issue; not only was the music non-diegetic, but the audiovisual experience itself was intermittent (cf. Figure 2). While subjects heard music throughout the experiment, they saw films for only 6 out of every 30 seconds; four-fifths of the experiment was devoted solely to listening to music. This may have precluded the music from binding to the visual domain, and consequently explain the mixed results seen in the present study. While subjective measures of valence and arousal suggest that music was able to impact the emotional response to viewing films, the physiology of response, though present, did not unfold as predicted. The affective information contained within the music thus may not have been believable at a level capable of influencing "gut" physiology, especially when paired with the affective information in the visual domain.

One potential solution to this binding problem would be to change the nature of the stimulus presentation; specifically, to present music *only* during the 6 s duration of each film. Such a design would be similar to that of Ellis and Simons (2002), in which a film and its associated soundtrack were exclusively presented together for the same amount of time (6 s). Such a short amount of time, however, may be insufficient for "non-content-loaded" music to interact with the "content-loaded" image. Such an experimental design, therefore, would almost necessarily preclude the use of music. Replacing musical stimuli with non-musical stimuli, however, might provide the answer. Bradley and Lang (2000a) suggested that short (6 s) nonmusical acoustic stimuli are effective in influencing both subjective and physiological indices of valence and arousal. The International Affective Digitized Sound system (IADS; Bradley & Lang, 1999), a library of some 115 acoustic stimuli, especially would be useful in the design of an audiovisual experiment of this nature.

Music and Emotion: Philosophical Perspectives

Perhaps one of the greatest difficulties in studying music and emotion is that there is considerable disagreement as to whether humans experience musical emotions in the same way they experience "gut" emotions (sadness, anger, fear, etc.). A person might say that they are "moved to tears" every time they listen to Barber's *Adagio for Strings*. Is this experience the same, however, as crying when emotionally distressed? There is no immediately clear answer. Levinson (1997) would deny that these two emotional responses are identical. A person cries when upset because something *made* that person upset—another individual, some piece of news, etc. That is, there is some object at which the emotional response is directed. Music, maintains Levinson, cannot function as an object; therefore the cause for our tears must have some other root.

The above situation is only one example from the larger debate on whether musical emotions are like nonmusical emotions. Kivy (1990) defined two positions on this issue: *emotivist* and *cognitivist*. The emotivist position assumes that, when we label a piece of music, say, "sad," it is because the music itself moves us to experience sadness (this might be akin to the notion of a "felt" emotion). The cognitivist position, on the other hand, assumes that when we label a piece of music sad it is because we recognize sadness as an expressive property of the music rather than actually being moved to sadness (akin to the notion of a "perceived" emotion). In other words, the emotivist position holds that music is expressive *of* emotion (rather like the face of a Saint Bernard is expressive *of* sadness) and that listeners recognize—but do not experience—this emotion. Alternatively, the cognitivist position maintains that listeners merely recognize emotions as expressive features of the music. Kivy (1990, p. 164) ultimately rejects the emotivist position, writing that music "like all of the arts, is . . . a thing of the intellect and not of the nerve endings."

By Way of a Conclusion: Music and the Human Experience

The present study suggests three points regarding music and film: First, that both filmic and musical stimuli elicit physiological changes; second, that an additive relationship exists between simultaneously presented music and film stimuli on some self report measures; and third, that the ability of music to bind to the visual narrative of a film may play an important role in enabling music to modulate the subsequent emotional response. We have not yet, however, discussed the question that in essence makes this research possible: *Why* is it that people believe in the existence of music when watching films? If one were to see an on-screen orchestra and hear orchestra-like sounds, it would be said that the music is part of the visual diegesis. But what about "true" film music? Where does it come from, and why do people accept its presence? Cohen (2001) argues that one might expect non-diegetic music to detract from (rather than add to) the sense of reality within a film. The results of this and other studies, however, suggest the opposite: pairing a film with emotionally-congruent music makes the experience *more* salient—subjectively across the board, and physiologically under circumstances that will require elaboration by further research.

References

Blood, A. J., & Zatorre, R. J. (2001). Intensely pleasurable responses to music correlate with activity in brain regions implicated in reward and emotion. *PNAS*, *98*(20), 11808–11823.

Bolivar, V. J., Cohen, A. J., & Fentress, J. C. (1994). Semantic and formal congruency in music and motion pictures: Effects on the interpretation of visual action. *Psychomusicology*, *13*(1 and 2), 28–59.

Boltz, M., Schulkind, M., & Kantras, S. (1991). Effects of background music on the remembering of filmed events. *Memory & Cognition*, *19*(6), 593–606.

Boltz, M. G. (2001). Musical soundtracks as a schematic influence on the cognitive processing of filmed events. *Music Perception*, *18*, 427–454.

Bradley, M. M., & Lang, P. J. (1999). *International affective digitized sounds (IADS): Stimuli, instruction manual, and affective ratings*. Technical Report B-2, The Center for Research in Psychophysiology, University of Florida.

Bradley, M. M., & Lang, P. J. (2000a). Affective reactions to acoustic stimuli. *Psychophysiology*, *37*, 204–215.

Bradley, M. M., & Lang, P. J. (2000b). Measuring emotion: behavior, feeling, and physiology. In R. D. Lane & L. Nadel (Eds.), *Cognitive Neuroscience of Emotion* (pp. 242–276). New York: Oxford.

Bullerjahn, C., & Güldenring, M. (1994). An empirical investigation of effects of film music using qualitative content analysis. *Psychomusicology*, *13*(1 and 2), 99–118.

Burt, G. (1994). *The art of film music*. Boston: Northeastern University Press.

Cacioppo, J. T., Berntson, G. G., Larsen, J. T., Poehlmann, K. M., & Ito, T. A. (2000). The psychophysiology of emotion. In M. Lewis, M., & J. M. Haviland-Jones (Eds.), *Handbook of Emotions* (2nd ed.) (pp.173–191). New York: Guilford Press.

Cohen, A. J. (1993). Associationism and musical soundtrack phenomena. *Contemporary Music Review*, *9*, 163–178.

Cohen, A. J. (2001). Music as a source of emotion in film. In P. N. Juslin & J. A. Sloboda (Eds.), *Music and Emotion: Theory and Research* (pp. 249–272). New York: Oxford.

Cook, E. W., III (1999). *VPM reference manual*. Birmingham, Alabama: Author.

Darwin, C. (1872/1969). *The Expression of the emotions in man and animals*. New York: Greeenwood Press.

Detenber, B. H. (1995). The effects of motion and image size on affective responses to and memory for pictures. Doctoral Dissertation, Stanford University, Stanford, CA.

Detenber, B. H., Simons, R. F., Roedema, T. M., & Reiss, J. E. (2000). The effects of color in film clips on emotional responses. *Media Psychology*, *2*, 331–355.

Deutsch, D., Ed. (1999). *The psychology of music*, 2nd ed. New York: Academic Press.

Dolan, R. J., Morris, J. S., & de Gelder, B. (2001). Crossmodal binding of fear in voice and face. *PNAS, 98*(17), 10006–10010.
Donnelly, K. J. (2001). The hidden heritage of film music: history and scholarship. In Donnelly (Ed.), *Film Music: Critical Approaches* (pp. 1–15). New York: Continuum.
Edwards, M. S., Eagle, C. T., Pennebaker, J. W., & Tunks, T. W. (1991). Relationships among elements of music and psychological responses of listeners. In C. D. Maranto (Ed.), *Applications of music in medicine* (pp. 41–57). Silver Spring, MD: National Association for Music Therapy.
Ellis, R. J., & Simons, R. F. (2002). The effects of sound on emotion during picture viewing. Poster presented at the 42nd Annual Meeting of the Society for Psychophysiological Research, Washington, DC.
Faith, M., and Thayer, J. F. (2001). A dynamical systems interpretation of a dimensional model of emotion. *Scandinavian Journal of Psychology, 42*, 121–133.
Gomez, P., Ott, R., & Danuser, B. (2002). Respiratory responses to environmental noises and music. Poster presented at the 42nd Annual Meeting of the Society for Psychophysiological Research, Washington, DC.
Izard, C. E. (1977). *Human emotions*. New York: Plenum.
Izard, C. E. (1993). Organizational and motivational functions of discrete emotions. In M. Lewis & J.M. Haviland (Eds.), *Handbook of Emotions* (pp. 631–642). New York: Guilford.
Juslin, P.N., & Sloboda, J.A. (2001). Music and emotion: Introduction. In Juslin & Sloboda (Eds.), *Music and emotion: theory and research* (pp. 3–20). New York: Oxford.
Juslin, P. N., & Zetner, M. R. (2002). Current trends in the study of music and emotion: Overture. *Music Scientiae, Special issue 2001–2002*, 3–21.
Kivy, P. *Music alone: Philosophical reflections on the purely musical experience.* Ithica: Cornell University Press, 1990.
Krumhansl, C. (1997). An exploratory study of musical emotions and psychophysiology. *Canadian Journal of Experimental Psychology, 51*(4), 336–352.
Lang, P. J. (1980). Behavioral treatment and bio-behavioral treatment: Computer applications. In J. B. Sidowski, J. H. Johnson, & T. A. Williams (Eds.), *Technology in mental health care delivery systems* (pp. 199–137). Norwood, NJ: Ablex.
Lang, P. J. (1995). The network model of emotion: Motivational connections. In R. S. Wyer & T. K. Srull (Eds.), *Advances in social cognition, 6,* Hillsdale, NJ: Erlbaum.
Lang, P. J., Greenwald, M. K, Bradley, M. M., & Hamm, A. O. (1993). Looking at pictures: Affective, facial, visceral, and behavioral reactions. *Psychophysiology, 30*, 261–273.
Lang, P. J., Öhman, A., & Vaitl, D. (1988). *The international affective picture system* [Photographic slides]. Gainesville: University of Florida, The Center for Research in Psychophysiology.
Lavy, M. M. (2001). Emotion and the experience of listening to music: A framework for empirical research. Ph.D. dissertation, University of Cambridge.
Levinson, J. (1997). Music and negative emotion. In J. Robinson (Ed.), *Music & meaning* (pp. 215–241). Ithaca: Cornell UP.
Lewis, M., & Haviland-Jones, J. M. (Eds.). (2000). *Handbook of Emotions* (2nd ed). New York: Guilford.
Neumann, S. A., & Waldstein, S. R. (2001). Similar patterns of cardiovascular response during emotional activation as a function of affective valence and arousal and gender. *Journal of Psychosomatic Research, 50*, 245–253.
Prendergast, R. M. (1992). *Film Music: A Neglected Art* (2nd ed.). New York: Norton.
Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology, 39*, 1161–78.
Russell, J. A. (1997). The circumplex in relation to emotions. In R. Plutchik & H. Conte (Eds.), *Circumplex models of personality and emotions* (pp. 205–221). Washington, D.C.: APA.

Russell, J. A. & Barrett, L.F. (1999). Core affect, prototypical emotional episodes, and other things called emotional: Dissecting the elephant. *Journal of Personality and Social Psychology*, *76*, 805-819.

Scherer, K. R., & Zentner, M. R. (2001). Emotional effects of music: production rules. In P. N. Juslin & J. A. Sloboda (Eds.), *Music and Emotion: Theory and Research* (pp. 361–392). New York: Oxford.

Simons, R. F., Detenber, B. H., Roedema, T. M. & Reiss, J. E. (1999). Emotion-Processing in Three Systems: The Medium and the Message. *Psychophysiology*, *36*, 619–627.

Sirius, G., & Clark, E. F. (1994). The perception of audiovisual relationships: A preliminary study. *Psychomusicology*, *13*(1 and 2), 119–132.

Sloboda, J. A., & Juslin, P. N. (2001). Psychological perspectives on music and emotion. In P. Juslin & J.A. Sloboda (Eds.), *Music and Emotion: Theory and Research* (pp. 71–104). New York: Oxford.

Tan, E. S. (1996). *Emotion and the structure of narrative film: Film as an emotion machine*. (B. Fasting, Trans.). Mahwah, NJ: Erlbaum.

Thayer, J. F., and Levenson, R. W. (1983). Effects of music on psychophysiological responses to a stressful film. *Psychomusicology*, *3*(1), 44–52.

Author note

Portions of this paper were first presented at the 43rd annual meeting of the Society for Psychophysiological Research, Chicago, IL, October, 2003, and were based on an undergraduate Honors thesis submitted to the University of Delaware by the first author.

Correspondence concerning this article should be addressed to Robert F. Simons, Department of Psychology, University of Delaware, Newark, DE, 19716. Electronic mail may be sent to either author: ellis.306@osu.edu or rsimons@udel.edu.

Appendix

Excerpts from the following eight works were used in the present study. They are grouped according to the four categories defined in the text.

Positive Valence, Low-Arousal

Beethoven, L.V. (1808). *Symphony 6 in F major, Op. 68*, 2nd mvmt. [Recorded by the Dresden Philharmonic, Herbert Kegel, director.] [CD]. Los Angeles: Laserlight 15 825.1988.

Vivaldi, A. (1725). *Guitar Concerto in D Major*, 2nd mvmt. [Recorded by the English Chamber Orchestra; Simon Wynberg, guitar; George Malcolm, cond.] On *In Classical Mood*, Vol. 11. [CD]. International Masters Publishers.1997.

Positive Valence, High-Arousal

Beethoven, L.V. (1805). *Symphony No. 3 in E-flat Major, Op. 55*, 3rd mvmt. [Recorded by the Hungarian Philharmonic Orchestra, Janos Ferencsik, cond.] [CD]. Laserlight 15 902. 1988.

Smetana, B. (1892). "Dance of the Comedians" from Act 3 of *The Bartered Bride*. [Performed by the London Symphony Orchestra, Geoffrey Simon, cond.] On *In Classical Mood*, Vol. 12. [CD]. International Masters Publishers.1997.

Negative Valence, Low-Arousal

Górecki, H. (1976). *Symphony No. 3, Op. 36*, 1st mvmt. [Performed by the Polish National Radio Symphony Orchestra, Andoni Wit, cond.] [CD]. On *Górecki: Symphony No. 3*. Munich: Naxos.1994.

Vaughan Williams, R. (1933). *A London Symphony*. 2nd mvmt. [Performed by the Philharmonia Orchestra, Owain Arwel Hughes, cond.] On *In Classical Mood*, Vol. 13. [CD]. International Masters Publiserhs.1997.

Negative Valence, High-Arousal

Schoenberg, A. (1930). *Begleitmusik zu einer Lichtspielszene*. [Incidental music to a motion picture scene.] [Recorded by the Chamber Orchestra of Europe, Heinz Holliger, cond.] [CD]. Berlin: Teldec 9031-77314-2. 1994.

Subotnik, M. (1985). *The Key to Songs*. [Performed by The California E.A.R. Unit.] [CD]. San Fransisco: Castro/New Albion Records.1986.